# Genome scan reveals new coat color loci in exotic pig cross

# Genome Scan Reveals New Coat Color Loci in Exotic Pig Cross

Hirooka H, de Koning DJ, van Arendonk JA, Harlizius B, de Groot PN, Bovenhuis H

## Abstract

The porcine genome was scanned to identify loci affecting coat color in an experimental cross between the Meishan breed and Dutch commercial lines. Linkage was studied in 1181 $F_2$ animals for 132 microsatellite markers and seven binary coat color scores: White, Black spotting, Speckle, Gray, Black, and specific color phenotypes for head and legs. The analyses were performed using interval mapping under various models. The study confirmed the existence of coat color loci on chromosome 8 and chromosome 6. One additional locus affecting White was detected on chromosome 5, possibly representing the porcine equivalent of the *steel* factor. Two new loci affecting Black were detected on chromosome 2. One of these showed exclusive maternal expression and mapped to a region where imprinted genes have been reported. The effect of the binary coding was tested by additional analyses excluding the white animals (>50% of $F_2$ animals). This showed that Black spotting was strongly influenced by the locus on chromosome 6 and the other color phenotypes were mainly influenced by the locus on chromosome 8. Epistatic effects were found between the loci on chromosomes 6 and 8 for Black spotting. For Black color, all combinations among chromosomes 2, 6, and 8 showed epistatic effects.

## Introduction

There are several varieties of coat color phenotypes in pigs. Most of the European breeds such as Large White and the different Landraces have a white coat color. On the other hand, the uniform black coat is the most common type in breeds from China (e.g., Legault 1998). Coat color inheritance in pigs has been extensively studied since the beginning of the last century. Inheritance of the white coat color was first studied bySpillman (1906) and he concluded that the white color is dominant. Hetzer (1945) confirmed this based on the results of crosses between Landrace and Large Black breeds. Recently Johansson et al. (1992) investigated an $F_2$ cross between the European wild boar and Large White pigs using genetic markers and showed that the gene for dominant white color was linked to the albumin (ALB) and the platelet-derived growth factor receptor α (*PDGFRA*) on chromosome 8. Johansson-Moller et al. (1996) also studied the white color locus using the same cross and showed that the alleles determining white coat color are associated with a duplication of the *KIT* gene. In pigs, four alleles have been reported at the *KIT* locus, the recessive *i* allele for color, the $I^P$ allele for black spots, the dominant *I*allele for the white phenotype, and $I^{Be}$ for the Belt phenotype (Giuffra et al. 1999; Johansson et al. 1992). Besides the *KIT* locus, strong evidence that the classical *Extension* coat color locus is identical to the melanocortin receptor 1 (*MC1R*) was found by Kijas et al. (1998). So far, four *MC1R*alleles have been reported: the $E^+$ allele for the wild type, the $E^d$ allele for the dominant black phenotype, the recessive *e* allele for the red color phenotype, and the $E^p$ allele for the spotted phenotype (black spotting). Results by Marklund et al. (1998) indicate that there are interactions between the *KIT* and the *MC1R* loci. Klungland and Våge (2000) have recently reviewed the different interactions that play a part in the pigmentation of domestic animals. Despite the important progress that has been made in recent years, the genetics of coat color variations has not been completely elaborated (Legault 1998).

Although coat color in pigs is not an economically important trait in the strict sense, consumers' preferences for pork from pigs with special coat colors still exists in some countries. For example, consumers in European countries prefer to eat pork from white-coated pigs. Therefore in these countries there is a strong preference for white coat color in commercial pig breeding. In contrast, Japanese consumers prefer to eat pork from black-coated pigs and such pork sells for a higher price.

The objective of this study was to detect, localize, and characterize effects of coat color loci in pigs using a whole genome scan based on more than 1000 $F_2$ animals from a cross between Meishan and Dutch pig breeds.

## Materials and Methods

### *Animals and Coat Color Phenotypes*

Phenotypic data for coat colors and marking patterns was collected on a cross ($F_1$ and $F_2$) of Chinese Meishan and Dutch pig lines originating from five different breeding companies. Meishan boars maintained at Wageningen University were crossed with Western sows at the different pig breeding companies to produce $F_1$ animals. A total of 1181 $F_2$ animals resulting from inter-se mating between $F_1$ boars and sows were used for linkage analysis. More details on the experimental design can be found inDe Koning et al. (1999) and Janss et al. (1997).
The color phenotype of newborn piglets was scored using a three-digit coding system. The first digit is for main color, where "main" is defined as more than 50%. The second digit is for shape and types of marks on the body. The most important shapes in our dataset are speckles (either gray or black) and black spots. Spots are black and have a gray border. The third digit indicates the presence of marks on the head or the legs, or a difference between the color of the head or the legs as compared to the rest of the body.

Derived from these codes, seven color phenotypes were defined: entirely white animals (White), entirely black animals (Black), animals with black spots (Black spotting), Speckles, and gray areas (Gray). For these phenotypes, sufficient numbers could be identified to include them in the analysis. Besides these nonoverlapping traits, we also defined animals with marks on the head or a different color of the head as compared to the color of the rest of the body (Head) and animals with marks on the legs or a different color of the legs as compared to the rest of body (Legs). The seven different coat color phenotypes are illustrated in Figure 1. For Head and Legs, different combinations are possible and therefore some typical examples are included in Figure 1. A description of the derived coat color phenotypes and their distribution in $F_1$ and $F_2$ individuals are shown in Table 1. For the analyses, the animals' coat colors were coded as 1 for expression and −1 for nonexpression for each of the described phenotypes.

*Genotyping and Statistical Analysis*

The $F_2$ animals, their $F_1$ parents, and the Meishan grandsires were typed for 132 microsatellite markers, which cover more than 90% of the porcine genome. DNA for genotyping was isolated from frozen blood or tissue. Genotyping of microsatellite markers was done as previously described (De Koning et al. 1999; Groenen et al. 1996). Details about polymerase chain reaction (PCR) mixtures, PCR conditions, and multiplexes can be found in Groenen et al. (1996). Multipoint recombination fractions were calculated using CriMap version 2.4 (Green et al. 1990). Recombination fractions were transformed to map distances using the Haldane mapping function. Interval mapping by line cross analysis was applied using regression methods (Haley et al. 1994). Based on the marker genotypes, probabilities were estimated for whether $F_2$ individuals inherited two Meishan alleles ($p_{11}$), two Dutch alleles ($p_{22}$), or one allele from each founder line ($p_{12}$ or $p_{21}$; different subscripts denoting parental origin; first subscript is paternally inherited allele). These probabilities were used in a least squares model to estimate the effect of a genome region on the coat color phenotype. The following one-locus model was fitted at every location on the genome:

$$y_i = \mu + ac_{a_i} + dc_{d_i} + e_i, \qquad (1)$$

where $\mu$ is the mean, $c_{ai}$ is the coefficient ($p_{11} - p_{22}$) for the additive component for individual $i$ at the given location and $c_{di}$ is the coefficient ($p_{12} + p_{21}$) for the dominant component for individual $i$ at the given location, $a$ and $d$ are the estimated additive and dominant effects of a putative locus at the given position, and $e_i$ is the residual error. The additive effect of the locus of interest is defined as half of the phenotypic difference between pigs homozygous for the locus from Meishan and Dutch alleles. A positive value denotes an increased expression of the coat color phenotype as a result of the Meishan alleles. The dominant effect is defined as the deviation of the heterozygous animals from the mean of the two types of homozygous animals.

Imprinting was tested following the procedures presented by De Koning et al. (2000). The contribution of the parents was separated using the probability that the individual inherited a Meishan allele from its father ($c_{\text{pat}} = [p_{11} + p_{12}] - [p_{22} + p_{21}]$) or from its mother ($c_{\text{mat}} = [p_{11} + p_{21}] - [p_{22} + p_{12}]$). This reparameterization allowed additional models to be fitted with exclusive paternal or maternal expression. All putative quantitative trait loci (QTL) locations

from the three models were subsequently evaluated with a saturated model that contained a paternal, maternal, and dominant component:

$$y_i = \mu + a_{pat}c_{pat_i} + a_{mat}c_{mat_i} + dc_{d_i} + e_i \qquad (2)$$

Using *F* ratios for the individual components of the model, imprinting was inferred if only one of the parental contributions was significant and no dominance was present. Further scrutiny for imprinting was obtained by an *F* test of the saturated model [equation (2)] versus the Mendelian model [equation (1)].

In addition, a two-way interaction model considering all epistatic effects was fitted to all combinations of significant, unlinked loci that were identified under the line cross models [equations (1) and (2)]:

$$\begin{aligned} y_i = {} & \mu + a_1c_{a1_i} + a_2c_{a2_i} + d_1c_{d1_i} + d_2c_{d2_i} \\ & + aac_{a1_i}c_{a2_i} + adc_{a1_i}c_{d2_i} + dac_{d1_i}c_{a2_i} \\ & + ddc_{d1_i}c_{d2_i} + e_i' \end{aligned} \qquad (3)$$

where variables are as in equation (1), with additional subscripts for locus 1 and 2, *aa* is the additive by additive epistatic interaction at the given locations, *ad* and *da* are the two dominant by additive interactions, and *dd*is the dominant by dominant interaction. First, the full models were fitted and subsequently nonsignificant components were dropped until the most parsimonious model remained.

To investigate segregation of QTL within families, QTL analyses were also performed under a paternal half-sib model (Knott et al. 1996). As opposed to the line cross models, the half-sib model makes no assumption about the number of QTL alleles and allele frequencies within the founder lines. For the half-sib analysis, the $F_2$ animals are treated as unrelated half-sib families and contrasts between the two haplotypes of every $F_1$ boar are studied. Half-sib families were only analyzed for a specific color phenotype if at least two animals within that family displayed the phenotype. The number of families included in the half-sib analysis varied between 10 for entirely black and gray, and 38 for entirely white. Within every half-sib family a QTL with a gene substitution effect is fitted at 1 cM intervals along the chromosome:

$$y_{ij} = f_j + b_j p_{ij} + e_{ij}, \quad (4)$$

where $y_{ij}$ is the trait score of individual $i$ originating from boar $j$, $f_j$ is the average effect for half-sib family $j$, $b_j$ is the substitution effect for a putative QTL, $p_{ij}$ is the conditional probability for individual $i$ of inheriting the first paternal haplotype, and $e_{ij}$ is the residual effect. The test statistic is calculated as an $F$ ratio for every map position within and across families. For details on half-sib analyses applied to this experimental population see De Koning et al. (1999). In the families that were inferred to be segregating for an identified QTL it was determined whether the Meishan allele was associated with an increase or a decrease in phenotype.

Significance thresholds for all models against the $H_0$ of no QTL were determined empirically using permutations (Churchill and Doerge 1994). A total of 10,000 permutations were conducted for each relevant chromosome. The genomewide threshold was obtained by accounting for testing on the other chromosomes using the Bonferoni correction (e.g., De Koning et al. 1999). To facilitate graphical comparisons of different models, a transformation was applied to the test statistics. The tabulated *P*value was calculated for every test statistic, using an $F$ distribution with the appropriate degrees of freedom. In the graphs the negative logarithm of these $P$ values is presented [$-\log_{10}(P)$].

## Results

### *The Distribution of Coat Color Phenotypes*

The distribution of coat color phenotypes for $F_1$ and $F_2$ animals is shown in Table 1. Of the 1181 $F_2$ individuals, 957 could be assigned to one of the nonoverlapping classes: White, Black spotting, Speckle, Gray, or Black. The main coat color phenotype in $F_1$ animals was White (54%). Among the $F_2$ animals, 616 of the animals were entirely white, and a wide diversity of coat color phenotypes including reddish brown (not shown) and gray were found. The mating of white $F_1$ animals produced 231 $F_2$ animals, of which 149 were white (65%) and 82 were colored (35%). This segregation pattern deviates from the expected 3:1 ratio under the assumption of a single dominant gene and heterozygosity of all white $F_1$ animals.

### *Line-Cross Analysis*

The genome scan showed genomewide significance on SSC2, SSC5, SSC6, and SSC8 (Table 2). Figures 2–4 show the profiles of the test statistics and the threshold levels along SSC8,

SSC6, and SSC2 for loci detected at 5% genomewide significance level. For White, the highest test statistic on SSC5 was between *IGF1* and SW378. The highest test statistic on SSC6 was at marker S0035 and on SSC8 at marker S0017. The putative loci affecting White on SSC6 accounted for 3% and on SSC8 for 36% of the $F_2$ variance. The areas affecting White on SSC6 and SSC8 were also significant for Black and Black spotting. For Speckle, Gray, Head, and Legs, only the locus on SSC8 showed genomewide significance. The loci on SSC8 affecting all coat color phenotypes were located between 75 and 82 cM, indicating that these phenotypes are probably influenced by a single locus.

A significant locus affecting Black was detected on SSC2, with the maximum peak near marker SWC9 (Figure 4 and Table 2). This QTL showed exclusive maternal expression with an $F$ ratio against a Mendelian model of 6.89 ($P < .01$) Furthermore, a smaller but also genomewide significant locus was observed on SSC2 at 129 cM near S0378. A two-dimensional search, fitting the coefficients for two locations on a chromosome simultaneously, was carried out to confirm the existence of two loci on SSC2 (Knott et al. 1998). The result indicated that including the second locus gave a significantly better fit ($P < .0001$) than the best single-locus model.

Table 2 also shows the additive and dominant effects for genomewide significant loci on SSC2, SSC5, SSC6, and SSC8. Under the standard line cross model the additive effects of the loci detected on both SSC6 and SSC8 were negative for the White phenotype, indicating that the proportion of white animals decreases with an increasing number of alleles from the Meishan breed. In contrast, the estimated additive and dominant effects for the locus on SSC5 are positive, indicating that a Meishan allele would result in a higher proportion of white animals. Since the phenotypes of coat colors are expressed by binary codes (−1 and 1), the additive effects are in the range of −1 and 1. For Black spotting, the additive effects of the loci on SSC6 and SSC8 were opposite. With respect to the loci on SSC8 for Speckle and Gray, only dominant effects were significant. For the locus on SSC8 affecting Head and Legs, both the additive and dominant effects were significant. The signs of dominant effects for Head and Legs were opposite. The effects of the loci detected on SSC2 for Black were 0.07 for the maternally expressed locus near SWC9, whereas the Mendelian locus around 129 cM had an estimated additive effect of −0.08 (Table 2). The opposite signs of these effects indicate that these two loci on SSC2 influence the expression of Black antagonistically.

If the suggestive linkage threshold is applied, 10 additional loci were detected on seven different chromosomes (Table 3). The region affecting White, Black spotting, and Black on

SSC6 is also significant at the suggestive level for Legs and Speckle. Two suggestive loci, affecting Gray and Black spotting, mapped to different regions on SSC9.

*Epistatic Interactions*

Table 4 shows the epistatic effects under the interaction models. For additive and dominant effects, the estimates from the interaction model, considering the epistatic effect, were comparable to the estimates of the single-locus model (not shown). For White there was significant dominant by dominant interaction between SSC5 and SSC6. For SSC6 and SSC8 there was significant additive (SSC6) by dominant (SSC8) interaction, as well as dominant by dominant interaction. For Black spotting, all possible epistatic interactions were significant between the loci on SSC6 and SSC8. Further, the maternally expressed locus on SSC2 showed significant interactions with the additive component of all three other loci affecting Black, and with the dominant component of the locus on SSC8. The second locus affecting Black on SSC2 showed additional additive by additive epistasis with SSC6 and SSC8 and dominant by additive as well as additive by dominant epistasis for SSC8. The loci on SSC6 and SSC8 show significant additive by additive interaction and additive (SSC6) by dominant (SSC8) interaction.

*Partial Data Analyses*

From the $F_2$ animals, 52% were entirely white. Given the binary structure of the derived phenotypes, this may affect the analyses for other coat colors. For any given non-white phenotype, the white animals are always included in the contrast. As a result, any identified loci may reflect a major effect on White rather than a true effect on the phenotype under study. To scrutinize whether some of the observed effects for the other colors were not merely caused by the effect of the locus on White, additional analyses (partial data analysis) were carried out after excluding all white animals. The results from the partial data analysis are shown in Table 5. For SSC6, the test statistic for the locus affecting Black spotting increased dramatically for the analysis of the partial data. In contrast, the test statistics for the locus affecting Black spotting on SSC8 and the locus affecting Black on SSC6 dropped below the genomewide threshold (Table 5). The test statistics for loci on SSC8 affecting Speckle and Gray increased, but the test statistics for loci affecting Black and Legs decreased. On SSC2, the test statistics of the imprinted locus at 3 cM affecting Black decreased slightly from 23.8 to 21.6, but remained highly significant. On the other hand, the test statistics of the

other putative locus (positioned at 129 cM in Table 3) on the same chromosome dropped below the genomewide significance threshold, but remained strongly suggestive (Table 5).

*Half-Sib Analyses*

The half-sib analysis showed genomewide evidence for a locus on SSC6 affecting White, and loci on SSC8 for White, Black spotting, Black, Head, and Legs. The estimated positions of the loci on SSC6 and SSC8 for White were identical to those detected by the line cross analysis. The estimated positions on SSC8 were 68 cM for Black spotting, 67 cM for Black, 66 cM for Head, and 75 cM for Legs. The test statistic for the individual families varied substantially. For the highly significant locus on SSC8 affecting White, four families had individual *F* ratios between 20 and 50, seven families had an *F* ratio between 10 and 20, and the remaining families had *F* ratios between 0 and 10. This clearly illustrates that even for this most significant locus, not all $F_1$ sires appear to be heterozygous. For the locus affecting Black on SSC8, only a single half-sib family showed an *F* ratio greater than 20, while all other families had *F* ratios less than 8. This indicates that the founder lines were not entirely fixed for the coat color loci, with the variation most likely coming from the commercial lines.

**Discussion**

The present study confirms that the most important coat color locus in pigs is located on chromosome 8. Johansson-Moller et al. (1996) pointed out that both the *I* and the $I^P$ alleles are associated with a duplication of the *KIT* gene. More recently, Marklund et al. (1998) found that the dominant white phenotype is caused by two mutations in the *KIT* gene: one gene duplication related with a partially dominant phenotype and a splice mutation in one of the copies leading to the fully dominant allele. The estimated position of the coat color locus on SSC8 detected in the present study is at 82 cM and this is close to the *KIT* locus. Therefore these results are consistent with results of the previous studies and it seems likely that the locus detected on SSC8 actually is the *KIT* locus. Giuffra et al. (1999) reported that the white belt gene in pigs was closely linked to marker S0086. Therefore the white belt gene might actually be an allele of the *KIT* locus. The white belt phenotype was not observed on any of the animals in this study.

The locus detected on SSC6 for White, Black, and Black spotting in the present study (Table 2) agrees with the *Extension* coat color locus (*E*locus, *MC1R*) reported by Kijas et al. (1998) and Mariani et al. (1996). Recently Kijas et al. (2001) came to the remarkable

conclusion that the black spots in $E^p/E^p$ individuals are actually the result of somatic reversions. Detecting a locus for the white color on SSC6 is not surprising because the opposite class of white color, namely "colored," includes various extension colors such as Black spotting, Speckle, and Black (see Materials and Methods). Therefore the locus on SSC6 resulted in an opposite effect on white coat color.

The detection of the *KIT* locus and the *MC1R* locus might not be surprising, as these loci are known to play an important role in pig coat color genetics (Johansson-Moller et al. 1996; Kijas et al. 1998, 2001;Mariani et al. 1996; Marklund et al. 1998). However, the interval mapping approach used in this study is usually applied to quantitative traits. The fact that this approach, in combination with the binary coding of the data, identified coat color loci that were described previously gives confidence in the method used.

The locus on SSC5 affecting white color has not been described previously. The locus maps to the interval IGF1–SW378. On the cytogenetic map (http://www.ri.bbsrc.ac.uk/cgi-bin/arkdb/rowsers/browser.sh?species = pig), both markers are assigned to 5q2.4-5q2.5, which includes*MYF5*. This region has been shown to be homologous to human chromosome 12q22 (http://www.toulouse.inra.fr/lgc/pig/compare/). This region includes the mast cell growth factor (*MGF*), also known as the steel factor, *kitl* (mice), or *KITLG* (human,http://www.cbc.umn.edu/tad/genes.htm). *MGF* is a ligand for the KIT receptor, and a variant of the gene causes, for example, the roan phenotype (Seitz et al. 1999) and white heifer disease in cattle (Charlier et al. 1996). So far *MGF* has been assigned to SSC5 by using a cell hybrid panel and CATS primers (http://www.ncbi.nlm.gov; accession number AF11990). We propose *MGF* as a positional candidate gene for the locus affecting White on SSC5.

The loci affecting black coat color on SSC2 have not been reported in any of the previous coat color studies in pigs. The first, and most significant, locus is located at a similar position as a paternally expressed QTL for muscle mass and fat deposition identified by Jeon et al. (1999) and Nezer et al. (1999). The tip of SSC2p is homologous to chromosome 11p15.5 in human (Rattink et al. 2001). No coat color or modifier loci mapping to that area in human or mice have been reported so far. The maternal expression for this locus might seem in contrast with the paternally expressed QTL identified in other studies (e.g., Hirooka et al. 2001; Jeon et al. 1999). However, in corresponding regions in human and mice a number of imprinted genes have been described, showing exclusive paternal or maternal expression (Morison et al. 2001).

Also the suggestive loci affecting coat color in pigs have not been reported in earlier studies (Table 3). However, since there is only suggestive evidence for these loci, further research will be required to confirm the effects of these loci.

For White, Black spotting, and Black phenotypes, interesting epistatic effects were found between the loci (Table 4). Marklund et al. (1998)showed that there are interactions between the *KIT* locus and the *MC1R*locus, which was confirmed by results in the present study. The importance of epistatic effects has been suggested from classical quantitative studies (e.g., Falconer 1981). Recently significant epistatic effects have also been revealed in QTL mapping studies (Wang et al. 1999). In interaction models, three types of epistasis have been considered: (1) between additive effects, (2) between additive and dominant effects, and (3) between dominant effects (Carlborg et al. 2000;Jana 1971). We fitted all these interactions to combinations of loci that were previously identified under Mendelian or imprinting models. Searching the genome with an interaction model might result in different estimates for the position of the loci and the magnitude of their effect (Carlborg et al. 2000). Also the validity of the full interaction model and the interpretation of the estimated effects for the binary coded phenotypes needs further investigation.

Our results indicate that the expression of black spotting and black colors may be caused by the epistatic effects between the *KIT* locus on SSC8, the*Extension* locus (*E* locus) on SSC6, and the new loci on SSC2. The biology of coat color would explain epistatic effects between *KIT, MGF* (ligand binding to *KIT* receptor), and the *E* locus (melanocortin 1 receptor) (Klungland and Våge 2000). Significant interactions were found between SSC5 (*MGF*) and SSC6 (*MC1R*), as well as and between SSC6 and SSC8 (*KIT*), but not between SSC5 and SSC8.

In order to provide insight into the implications of the model used in the present study, Table 6 shows the predicted effects of different Meishan and Dutch allele combinations for the genes detected on chromosome 6 and 8, that is, presumably the *Extension* locus and the *KIT* locus, respectively. Predicted effects were calculated using only the significant effects in model 3, including significant effects due to interactions. Effects are reported for the White, Black, and Black spotting phenotypes, that is, the phenotypes showing interactions (Table 4). For White there is a gradually decreasing expression of the phenotype when moving from two Dutch alleles at both loci (0.96) to two Meishan alleles at both loci (−1.02). The effect of the *Extension* locus is limited in case the *KIT* locus is homozygous (MM or DD). However, if the *KIT* locus is heterozygous (MD), the effect of the *Extension* locus is more pronounced. There is no expression (−1.00) of the Black phenotype if two Dutch alleles

have been inherited at the *KIT* locus. An incomplete expression of Black is observed if two Meishan alleles have been inherited both at the *KIT* locus and the*Extension* locus. The value of −0.13 for this genotype combination indicates that not all of the animals that inherited only Meishan alleles at the *KIT* and the *Extension* loci are completely Black (complete expression would imply a value of 1). A value of −1.02 for White further suggests that there are no White animals in this category. Consequently there will be several animals with marks. This illustrates that in order to inherit the Meishan coat color phenotype (entirely black), it is not sufficient to have Meishan alleles at the *KIT* and *Extension* loci. Inspection of the coat color codes for animals with a probability of greater than 80% of having Meishan alleles at both the *KIT* and the *Extension* loci shows that almost all of them have black as the main color (more than 50% of the body is black). However, the number of entirely black animals is limited, as several have white marks, especially on the head and/or the legs. Similar as for Black, Black spotting is only partly expressed. This is the case if the *KIT*locus has genotype MM and the *Extension* locus has genotype DD. Note that this pattern of almost no expression for all genotypes except for one results in significant AA, AD, DA, and DD interactions.

Besides loci on SCC6 and SCC8, Black is also influenced by two loci on SCC2. Several epistatic interactions between these loci significantly affect the Black phenotype. Model 3 would predict that the most extremely black-colored animal would carry a maternally inherited Meishan allele at SSC2 (3 cM), two Dutch alleles at SSC2 (129 cM), two Meishan alleles at SSC6, and two Meishan alleles at SSC8. Remarkable is that at one of the loci on SSC2 (129 cM) the Dutch alleles increase the Black phenotype. The Dutch alleles at this locus that increase the Black phenotype are recessive to the Meishan allele (see Table 2).

Results of Kijas et al. (1998) and Marklund et al. (1998) suggest that the genotype of Meishan is *i*/*i* for the *KIT* locus and $ED^1/ED^1$ for the *MC1R*locus. For the Dutch lines it is not clear which breeds they comprise, but Dutch Landrace and Large White are important contributors to the Dutch commercial lines. Based on the results of Kijas et al. (1998) and Marklund et al. (1998) it can be hypothesized that genotypes are $E^p/E^p$ for the *MC1R*locus and *I*/- for the *KIT* locus (*I*/*I*, $I/I^p$, or *I*/*i*). The possible segregation at the *KIT* locus hinders the interpretation of the effects reported in Table 6. It seems reasonable, however, to assume that the frequency of $I^p$ and *i*alleles in the Dutch lines is low, as breeding organizations prefer a uniform product and selection will be directed toward entirely white pigs. Therefore we can assume that the genotype MM for *KIT* and DD for*Extension* correspond to the *i*/*i*, $E^p/E^p$ genotype. According to Marklund et al. (1998), this genotype would result in pigs that are white with black spots or red with black spots.

Inspection of the coat color codes for animals with a probability of greater than 80% of having Meishan alleles at the *KIT* locus and Dutch alleles at the *Extension* locus demonstrates that approximately half of them have black spots. However, also entirely white and entirely black animals do occur.

The coding of color phenotypes into a binary format makes it easier to contrast between one specific phenotype and all the others. This implies that it is difficult to determine which side of the phenotypic contrast is affected by an identified locus. In this study this was apparent for the white color, with 52% of the $F_2$ animals being entirely white. Therefore we tried to redefine the contrasts for the nonwhite color phenotypes by removing data for animals with white coat color (i.e., partial data analysis). The results indicated that the locus on SSC6 might have a stronger effect on black spotting than the locus on SSC8, although the locus on SSC8 was more significant than that on SSC6 under the analysis using all data (Table 2 versus Table 5). This illustrates that there are dependencies among coded phenotypes, partly as a result of the binary trait definitions. The partial analysis, excluding animals displaying the most prevalent phenotype, might provide some insight into these dependencies, but differences in sample size between full and partial data complicate the final conclusions.

Using the three initial codes adopted for classification of coat colors in this study, it is possible to define other coat color phenotypes, for example, basic white color (>50% white color), basic black color (>50% black color), and white leg color combined with black as the main color. When using these color phenotypes, genomewide significant loci were detected on SSC8 for all color phenotypes, on SSC6 for basic white and black colors, and on SSC2 for basic black color. The best positions of putative loci were comparable to those reported in Table 2. This illustrates the interrelationships among coat color phenotypes, because categories of these phenotypes overlap. Furthermore, these results show that the linkage analysis for coat color loci is fairly robust against the definition of the binary trait scores and the major effect of SSC8 on the expression of various coat colors.

In this study the coat color phenotypes were treated as continuous variables, although they were scored as binary variables (i.e., −1 and 1). In general, methods for mapping QTL based on regression methods assume that residuals are normally distributed. However, this assumption is not crucial when thresholds are determined empirically. Theoretically, analyses using the binary data may require a special approach. Jansen (1992) presented a general mixture model for mapping QTL, which fits a generalized linear model (GLM). However, using stochastic simulation,Visscher et al. (1996) showed that regression methods can be

used for mapping QTL for binary data. In practice, regression methods have the advantages that they are easy and quick to use.

For most analyses in this study we adopted the line cross approach, that is, we contrasted Meishan and Dutch Large White alleles, assuming both lines to be fixed for alternative alleles for coat color loci. However, alleles are not necessarily fixed in the Meishan and/or the Dutch population, as was illustrated by the results from the half-sib analyses. The Dutch commercial lines are expected to be genetically heterogeneous. Alfonso and Haley (1998) pointed out that when the populations that are crossed are not fixed for alternative alleles, the power to detect a locus will be reduced and its effect will be underestimated as the locus allele frequencies in the two populations become more similar. If the marker is fixed at alternative alleles but the gene of interest is still segregating in both lines (i.e., outbred populations), the estimated regression coefficient represents $|p_1 - p_2|\,a$ instead of $a$, where $p_1$ and $p_2$ are the allele frequencies of line 1 and 2, respectively, and $a$ is the true additive effect. Similarly, the estimated effects for the interaction components are expected to be biased if alleles are segregating in both lines.

Although coat color at first sight seems to be a simple trait, the underlying genetics are more complicated. In pigs, coat color is influenced by one locus with a major effect (*KIT*), but there also appear to be a number of loci with smaller effects that interact not only among themselves, but with the *KIT* locus. Our findings indicate that all color phenotypes except Black spotting are strongly influenced by the *KIT* locus on SSC8, whereas Black spotting is influenced by the *Extension* coat color locus on SSC6. Furthermore, two loci on SSC2, which affect antagonistically the expression of black color, were detected at a 5% genomewide significance level. The loci on SSC2 and SSC5 have not been reported in previous studies, but for SSC5 there is a strong positional candidate. In addition to these loci, several new suggestive loci were detected on different chromosomes and at different positions. These results could help researchers identify additional markers and enhance more advanced studies on coat color in pigs.

*Acknowledgments*

This research was supported financially by the Netherlands Technology Foundation (to S.T.W.). Additional financial support was provided by the Dutch Product Board for Livestock, Meat, and Eggs and the Dutch pig breeding organizations Hypor BV, Dumeco Breeding BV, and Topigs. H. Hirooka acknowledges the financial support of The Netherlands' National Research Council (NWO) and Japan Society for the Promotion of

Science (JSPS). The European Union provided financial support to B. Harlizius. H. Hirooka and D. J. de Koning contributed equally to this article.

**References**


- Alfonso L and Haley CS, 1998. Power of different F2 schemes for QTL detection in livestock. Anim Sci 66:1–8.
- Carlborg O, Andersson L, and Kinghorn B, 2000. The use of a genetic algorithm for simultaneous mapping of multiple interacting quantitative trait loci. Genetics 155:2003–2010.
- Charlier C, Denys B, Coppieters W, Grobet L, Mni M, Womack J, Hanset R, and Goerges M, 1996. Microsatellite mapping of the bovine roan locus: a major determinant of White Heifer disease. Mamm Genome 7:138–142.
- Churchill GA and Doerge RW, 1994. Empirical threshold values for quantitative trait mapping. Genetics 138:963–971.
- De Koning DJ, Janss LGL, Rattink AP, van Oers PAM, de Vries BJ, Groenen MAM, van der Poel JJ, de Groot PN, Brascamp EW, and van Arendonk JAM, 1999. Detection of quantitative trait loci for backfat thickness and intramuscular fat content in pigs (Sus scrofa). Genetics 152:1679–1690.
- De Koning DJ, Rattink AP, Harlizius B, van Arendonk JAM, Brascamp EW, and Groenen MAM, 2000. Genome-wide scan for body composition in pigs reveals important role of imprinting. Proc Natl Acad Sci USA 97:7947–7950.
- Falconer DS. 1981. Introduction to quantitative genetics, 2nd ed. New York: Longman Press.
- Giuffra E, Evans G, Tornsten A, Wales R, Day A, Looft H, Plastow G, and Andersson L, 1999. The Belt mutation in pigs is an allele at the dominant white (I/KIT) locus. Mamm Genome 10:1132–1136.
- Green P, Fallis K, and Crooks S. 1990. Documentation for CRIMAP version 2.4.
- Groenen MAM, de Vries BJ, van der Poel JJ. 1996. Alignment of the PiGMap and USDA linkage maps of porcine chromosomes 3 and 9. Anim Genet 27:355–357.
- Haley CS, Knott SA, and Elsen JM, 1994. Mapping quantitative trait loci in crosses between outbred lines using least squares. Genetics 136:1195–1207.
- Hetzer HO, 1945. Inheritance of coat color in swine. IV. Analysis of hybrids of Landrace and Large Black. J Hered 36:309–312.

- Hirooka H, de Koning DJ, Harlizius B, van Arendonk JAM, Rattink AP, Groenen MAM, Brascamp EW, and Bovenhuis H, 2001. A whole genome scan for quantitative trait loci (QTL) affecting teat number in pigs. J Anim Sci 79:2320–2326.
- Jana S, 1971. Simulation of quantitative characters from quantitatively acting genes. I. Non-allelic gene interactions involving two or three loci. Theor Appl Genet 41:216–226.
- Jansen RC, 1992. A general mixture model for mapping quantitative trait loci by using molecular markers. Theor Appl Genet 85:252–260.
- Janss LLG, van Arendonk JAM, and Brascamp EW, 1997. Baysian statistical analyses for presence of single genes affecting meat quality traits in a crossbred pig population. Genetics 145:395–408.
- Jeon J-T, Carlborg O, Tornsten A, Griuffra E, Amarger V, Chardon P, Andersson-Eklund L, Andersson K, Hansson I, Lundstron K, and Andersson L, 1999. A paternally expressed QTL affecting skeletal and cardiac muscle mass in pig maps to the IGF2 locus. Nat Genet 21:157–158.
- Johansson M, Ellegren H, Marklund L, Gustarsson U, Ringmar-Cederberg E, Andersson K, Edfors-Lilja I, and Andersson L, 1992. The gene of dominant white color is closely linked to ALB and PDGFRA on chromosome 8. Genomics 14:965–969.
- Johansson-Moller M, Chandhary R, Hellman E, Hoyheim B, Chowdhary B, and Andersson L, 1996. Pigs with the dominant white coat color phenotype carry a duplication of the KIT gene encoding the mast/stem cell growth factor receptor. Mamm Genome 7:822–830.
- Kijas JMH, Moller M, Plastow G, and Andersson L, 2001. A frameshift mutation in MC1R and a high frequency of somatic reversions cause black spotting in pigs. Genetics 158:779–785.
- Kijas JMH, Wales R, Tornsten A, Chardon P, Moller M, and Andersson L, 1998. Melanocortin receptor 1 (MC1R) mutations and coat color in pigs. Genetics 150:1177–1185.
- Klungland H and Våge DI, 2000. Molecular genetics of pigmentation in domestic animals. Curr Genomic 1:223–242.
- Knott SA, Marklund L, Haley CS, Andersson K, Davies W, Ellegren H, Fredholm M, Hansson I, Hoyheim B, Lundstrom K, Moller M, and Andersson L, 1998. Multiple

marker mapping of quantitative trait loci in a cross between outbred wild boar and Large White pigs. Genetics 149:1069–1080.

- Knott, SA, Elsen, JM, and Haley, CS, 1996. Methods for multiple-marker mapping of quantitative trait loci in half-sib populations. Theor Appl Genet 93:71–80.
- Lander ES and Kruglyak L, 1995. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241–247.
- Legault C, 1998. Genetics of colour variation: In: The genetics of the pig (Rothschild MF and Ruvinsky A, eds). New York: CAB International; 51–70.
- Mariani P, Johansson-Moller M, Hoyheim B, Marklund L, Davies W, Ellegren H, and Andersson L, 1996. The extension coat color locus and the loci for blood group O and tyrosine aminotransferase are on pig chromosome 6. J Hered 87:272–276.
- Marklund S, Kijas J, Rodriguez-Martinez H, Ronnstrand L, Funa K, Moller M, Lange D, Edfors-Lilja I, and Andersson L. 1998. Molecular basis for the dominant white phenotype in the domestic pig. Genome Res 8:826–833.
- Morison IM, Paton CJ, and Cleverley SD, 2001. The imprinted gene and parent-of-origin effect database. Nucleic Acids Res 29:275–276.
- Nezer C, Morean L, Brouwers B, Coppieters W, Detilleux J, Hanset R, Karim L, Krasz A, Leroy P, and Georges M, 1999. An imprinted QTL with major effect on muscle mass and fat deposition maps to the IGF2 locus n pigs. Nat Genet 21:155–156.
- Rattink AP, Faivre M, Jungerius BJ, Groenen MAM, and Harlizius B, 2001. A high-resolution comparative RH map of porcine chromosome (SSC) 2. Mamm Genome 12:366–370.
- Seitz JJ, Schmutz SM, Thue TD, and Buchanan FC, 1999. A missense mutation in the bovine MGF gene is associated with the roan phenotype in Belgian Blue and Shorthorn cattle. Mamm Genome 10:710–712.
- Spillman WJ, 1906. Inheritance of coat colour in swine. Science 24:441–443.
- Visscher PM, Haley CS, and Knott SA, 1996. Mapping QTL for binary traits in backcross and F2 populations. Genet Res 68:55–63.
- Wang DL, Zhu J, Li ZK, and Paterson AH, 1999. Mapping QTLs with epistatic effects and QTL × environment interactions by mixed linear model approaches. Theor Appl Genet 99:1255–1264.

**Legends**

Figure 1 Typical illustrations of the seven different coat color phenotypes: 1, White; 2, Black spotting; 3, Speckle; 4, Gray; 5, Black; 6, Head; 7, Legs**.**

Figure 2. Test statistic profile for all color phenotypes on chromosome 8.

Figure 3. Test statistic profile for the chromosome 6 with regard to White, Black spotting, and Black. The horizontal line denotes the 5% genomewide threshold.

Figure 4. Test statistic profile for chromosome 2 with regard to black phenotype under Mendelian and exclusive maternal and paternal expression. The horizontal line denotes the 5% genomewide threshold.

Table 1. Abbreviations, description of phenotype, and number of animals displaying the phenotype in a total of 291 $F_1$ and 1181 $F_2$ animals.

Table 2. Most likely positions, test statistics, and estimated effects for genomewide significant QTL affecting coat color phenotypes.

Table 3. Lists of suggestive QTL affecting coat color phenotypes.

Table 4. Significant epistatic effects between identified loci after elimination of nonsignificant components of equation (2)

Table 5. Most likely positions and test statistics for QTL affecting coat color phenotypes in partial data analysis[a]

Table 6. Predicted effects of the loci on SSC6 and SSC8 for White, Black, and Black spotting coat color phenotypes.

Fig.1

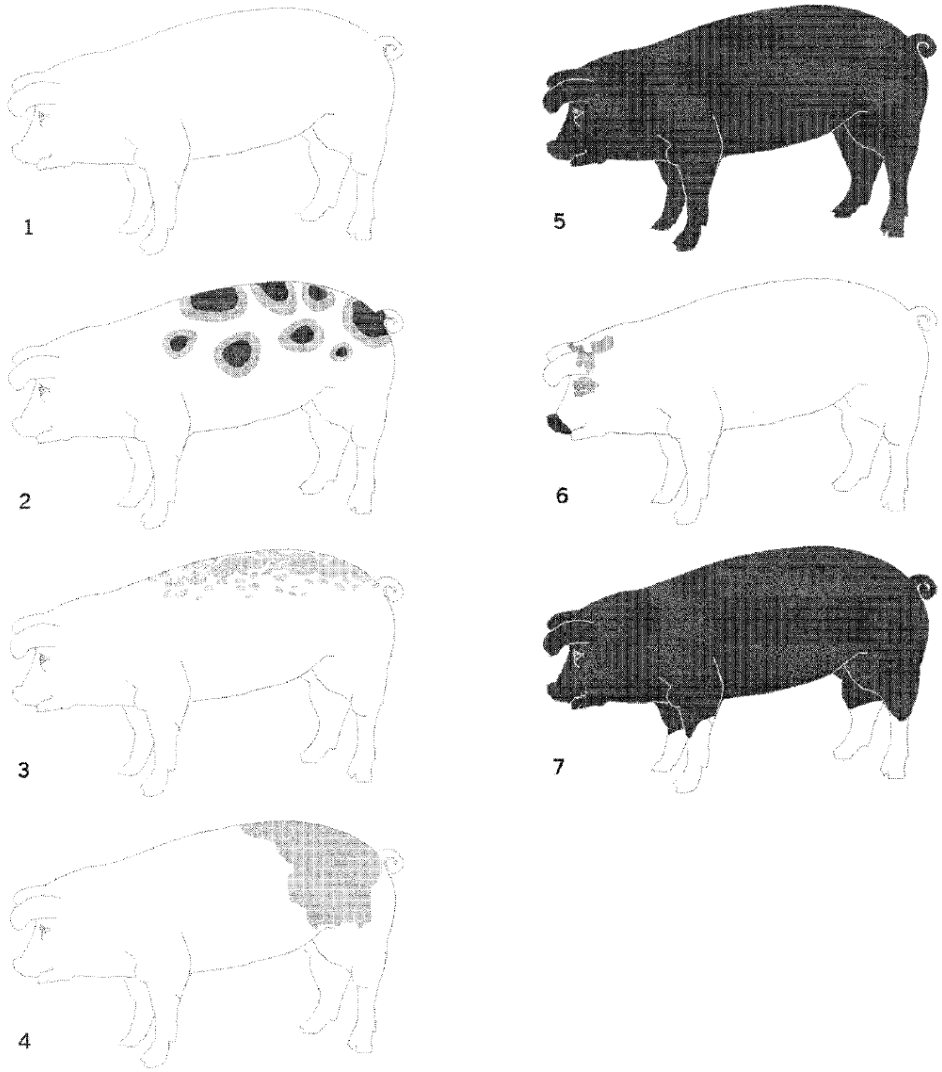


Fig.2

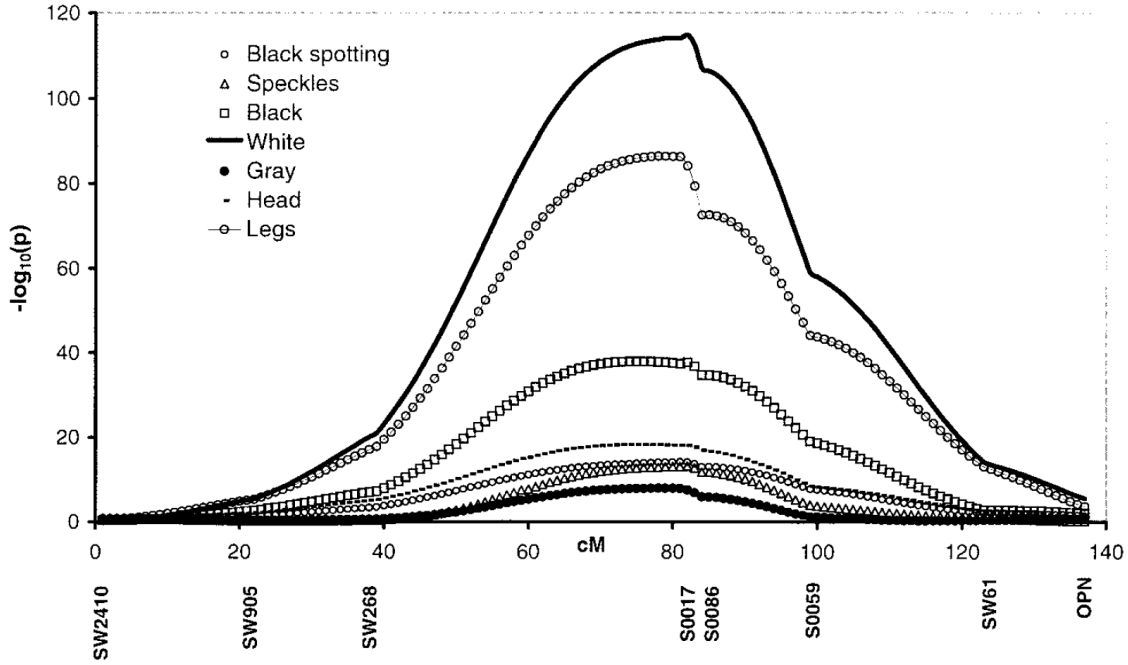

Black spotting
Speckles
Black
White
Gray
Head
Legs
-log10(p)
cM
SW2410
SW905
SW268
S0017
S0086
S0059
SW61
OPN

Fig.3

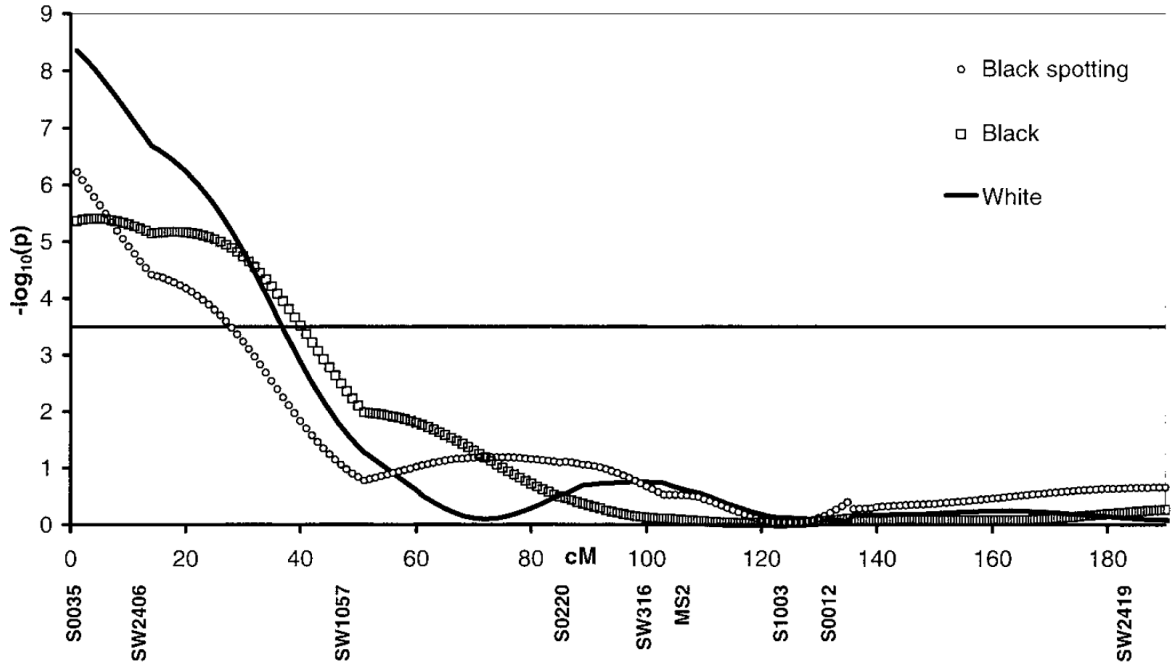
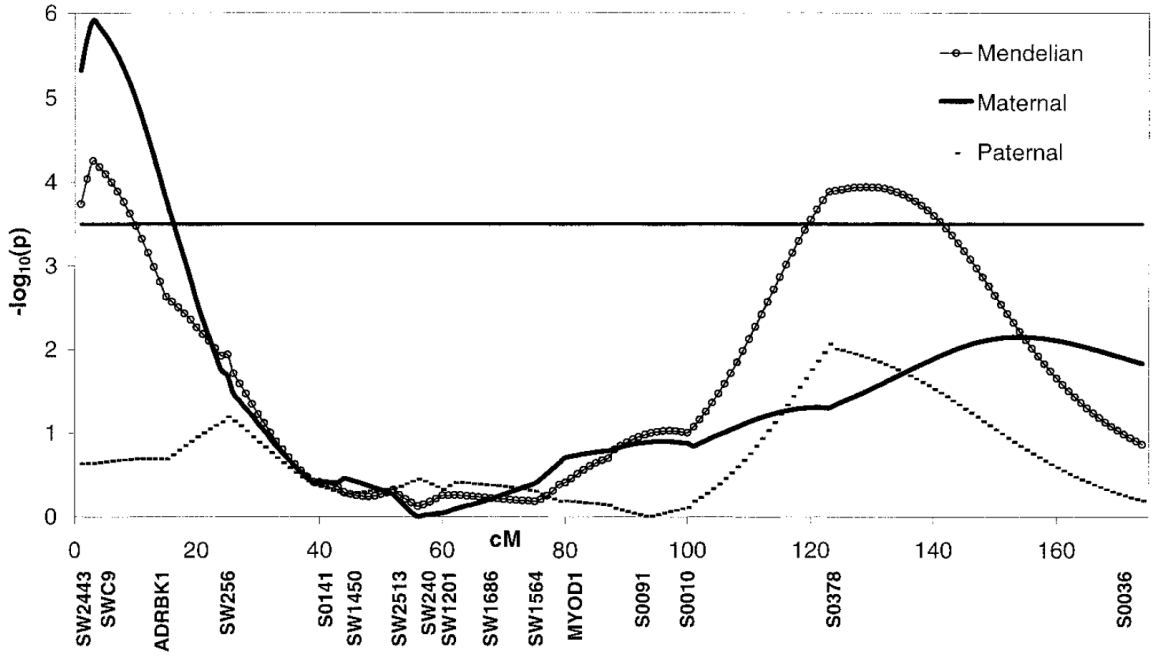

Black spotting
Black
White
-log10(p)
cM
S0035
SW2406
SW1057
S0220
SW316
MS2
S1003
S0012
SW2419

Fig.4


Mendelian
Maternal
Paternal
-log10(p)
cM
0
20
40
60
80
100
120
140
160
1
2
3
4
5
6
SW2443
SWC9
ADRBK1
SW256
S0141
SW1450
SW2513
SW240
SW1201
SW1686
SW1564
MYOD1
S0091
S0010
S0378
S0036

Tab.1

| Trait | Description and details | No. of measurements | |
|---|---|---|---|
| | | $F_1$ | $F_2$ |
| White | Entirely white color, no marks | 153 | 616 |
| Black spotting | Black marks with a gray border | 14 | 79 |
| Speckle | Speckle marks on the body | 91 | 133 |
| Gray | Gray area on the body | 22 | 63 |
| Black | Entirely black color, no marks | 0 | 66 |
| Head | Marks on the head or different color of the head as compared to the rest of the body | 74 | 285 |
| Legs | Marks on the legs or different color of the legs as compared to the rest of the body | 8 | 228 |

Tab.2

| Color phenotype | SSC | Position (cM) | Test statistic | Genomewide *P* value | Additive effect | Dominant effect |
|---|---|---|---|---|---|---|
| White | 5 | 108 | 7.80 | .05 | 0.21 ± 0.07 | 0.33 ± 0.14 |
| | 6 | 1 | 19.56 | <.001 | −0.34 ± 0.05 | −0.03 ± 0.10 |
| | 8 | 82 | 334.00 | <.001 | −0.90 ± 0.04 | 0.16 ± 0.05 |
| Black spotting | 6 | 1 | 17.73 | <.001 | −0.12 ± 0.03 | −0.15 ± 0.05 |
| | 8 | 81 | 33.36 | <.001 | 0.17 ± 0.02 | −0.07 ± 0.03 |
| Speckles | 8 | 82 | 31.04 | <.001 | 0.00 ± 0.03 | 0.31 ± 0.04 |
| Gray | 8 | 78 | 18.43 | <.001 | 0.00 ± 0.02 | 0.20 ± 0.03 |
| Black | 2[a] | 3 | 23.78 | <.001 | 0.07 ± 0.02 | NA |
| | 2 | 129 | 9.14 | .022 | −0.08 ± 0.03 | −0.12 ± 0.04 |
| | 6 | 4 | 12.59 | <.001 | 0.12 ± 0.02 | 0.00 ± 0.04 |
| | 8 | 75 | 94.20 | <.001 | 0.24 ± 0.02 | −0.27 ± 0.03 |
| Head | 8 | 75 | 43.61 | <.001 | 0.36 ± 0.04 | 0.21 ± 0.07 |
| Legs | 8 | 78 | 236.80 | <.001 | 0.61 ± 0.03 | −0.45 ± 0.05 |

[a] Exclusive maternal expression.

Tab.3

| Color phenotype | SSC | Position (cM) | Test statistic |
|---|---|---|---|
| White | 1 | 81 | 6.10 |
| Black spotting | 3 | 18 | 5.16 |
| | 4 | 78 | 5.29 |
| | 5[a] | 2 | 10.20 |
| | 9 | 58 | 6.38 |
| Speckle | 6 | 6 | 6.32 |
| | 13 | 32 | 5.93 |
| Gray | 9 | 96 | 6.99 |
| Head | 6 | 89 | 5.63 |
| Legs | 6 | 1 | 5.07 |

[a] Exclusive paternal expression, $F$ ratio against Mendelian model 3.65 ($P = .056$).

Tab.4

| | | | Type of epistasis[a] | | | |
|---|---|---|---|---|---|---|
| Color | SSC | SSC | AA | AD | DA | DD |
| White | 5 | 6 | — | — | — | −0.92 ± 0.44* |
| | 6 | 8 | — | −0.40 ± 0.09*** | — | −0.26 ± 0.11* |
| Black spotting | 6 | 8 | −0.29 ± 0.04*** | 0.33 ± 0.05*** | −0.31 ± 0.07*** | 0.32 ± 0.09*** |
| Black | 2[b] | 2[c] | −0.07 ± 0.03* | — | NA | NA |
| | 2[b] | 6 | 0.07 ± 0.03* | — | NA | NA |
| | 2[b] | 8 | 0.17 ± 0.02*** | −0.17 ± 0.04*** | NA | NA |
| | 2[c] | 6 | −0.10 ± 0.03* | — | — | — |
| | 2[c] | 8 | −0.11 ± 0.04** | 0.13 ± 0.06* | −0.16 ± 0.07* | — |
| | 6 | 8 | 0.19 ± 0.04*** | −0.23 ± 0.06*** | — | — |

*$P < .05$, **$P < .01$, ***$P < .001$.

[a] AA = additive × additive, AD/DA = additive × dominant, DD = dominant × dominant.

[b] Maternally expressed locus at 3 cM, used $c_{mat}$ instead of $c_{\sigma^2}$, no dominance interactions from this locus.

[c] Mendelian locus at 129 cM.

Tab.5

| Color phenotype | SSC (cM) | Position | Test statistic | Genome-wide *P* value |
|---|---|---|---|---|
| Black spotting | 6 | 1 | 35.54 | <.001 |
| | 8 | 61 | 4.48 | >.10 |
| Speckle | 8 | 82 | 64.45 | <.001 |
| Gray area | 8 | 78 | 29.05 | <.001 |
| Black | 2[b] | 3 | 21.61 | <.001 |
| | 2 | 127 | 7.37 | .12 |
| | 6 | 7 | 7.09 | .11 |
| | 8 | 73 | 29.57 | <.001 |
| Head | 8 | 81 | 18.09 | <.001 |
| Leg | 8 | 79 | 46.62 | <.001 |

[a] Entirely white $F_2$ animals were excluded from the analysis.

[b] Exclusive maternal expression.

Tab.6

| SSC8 | | SSC6 MM[a] | MD | DD |
|---|---|---|---|---|
| MM | White | −1.02 | −0.93 | −0.84 |
| | Black | −0.13 | −0.54 | −0.94 |
| | Black spotting | −0.92 | −0.97 | 0.23 |
| MD | White | −0.24 | −0.01 | 0.74 |
| | Black | −1.03 | −1.02 | −1.01 |
| | Black spotting | −0.86 | −0.90 | −0.95 |
| DD | White | 0.78 | 0.87 | 0.96 |
| | Black | −0.96 | −0.99 | −1.02 |
| | Black spotting | −1.00 | −1.00 | −1.00 |

[a] MM: two Meishan alleles; MD: one Meishan and one Dutch allele; DD: two Dutch alleles.